

# Capacities of classical compound quantum wiretap and classical quantum compound wiretap channels

Minglai Cai  
Department of Mathematics  
University of Bielefeld  
Bielefeld, Germany  
Email: mlcai@math.uni-bielefeld.de

Ning Cai  
The State Key Laboratory of  
Integrated Services Networks  
University of Xidian  
Xian, China  
Email: caining@mail.xidian.edu.cn

Christian Deppe  
Department of Mathematics  
University of Bielefeld  
Bielefeld, Germany  
Email: cdeppe@math.uni-bielefeld.de

**Abstract**—We determine the capacity of the classical compound quantum wiretapper channel with channel state information at the transmitter. Moreover we derive a lower bound on the capacity of this channel without channel state information and determine the capacity of the classical quantum compound wiretap channel with channel state information at the transmitter.

## I. INTRODUCTION

The compound channel models transmission over a channel that may take a number of states, its capacity was determined by [6]. A compound channel with an eavesdropper is called a compound wiretap channel. It is defined as a family of pairs of channels  $\{(W_t, V_t) : t = 1, \dots, T\}$  with common input alphabet and possibly different output alphabets, connecting a sender with two receivers, one legal and one wiretapper, where  $t$  is called a state of the channel pair  $(W_t, V_t)$ . The legitimate receiver accesses the output of the first channel  $W_t$  in the pair  $(W_t, V_t)$ , and the wiretapper observes the output of the second part  $V_t$  in the pair  $(W_t, V_t)$ , respectively, when a state  $t$  governs the channel. A code for the channel conveys information to the legal receiver such that the wiretapper knows nothing about the transmitted information. This is a generalization of Wyner's wiretap channel [15] to the case of multiple channel states.

We will be dealing with two communication scenarios. In the first one only the transmitter is informed about the index  $t$  (channel state information (CSI) at the transmitter), while in the second, the legitimate users have no information about that index at all (no CSI).

The compound wiretap channels were recently introduced in [9]. A upper bound on the capacity under the condition that the average error goes to zero and the sender has no knowledge about CSI is obtained. The result of [9] was improved in [5] by using the stronger condition that the maximal error should go to zero. Furthermore, the secrecy capacity for the case with CSI was calculated.

This paper is organized as follows.

In Section II we present some known results for classical compound wiretap channel which we will use for our result's proof.

In Section III we derive the capacity of the classical compound quantum wiretap channel with CSI and give a lower

bound of the capacity without CSI. In this channel model the wiretapper uses classical quantum channels.

In Section IV we derive the capacity of the classical quantum compound wiretap channel with CSI. In this model both the receiver and the wiretapper use classical quantum channels, and the set of the states can be both finite or infinite. Here we will use an idea which is similar to the one used in [2].

## II. CLASSICAL COMPOUND WIRETAP CHANNELS

Let  $A, B$ , and  $C$  be finite sets,  $P(A)$ ,  $P(B)$ , and  $P(C)$  be the sets of probability distributions on  $A$ ,  $B$  and  $C$ , respectively. Let  $\theta := \{1, \dots, T\}$ . For every  $t \in \theta$  let  $W_t$  be a channel  $A \rightarrow P(B)$  and  $V_t$  be a channel  $A \rightarrow P(C)$ . We call  $(V_t, W_t)_{t \in \theta}$  a compound wiretap channel.  $W_t^n$  and  $V_t^n$  stand for the  $n$ -th memoryless extensions of stochastic matrices  $W_t$  and  $V_t$ .

Here the first family represents the communication link to the legitimate receiver while the output of the latter is under control of the wiretapper.

Let  $X$  be a discrete random variable on a finite set  $\{x_1, \dots, x_n\}$ , with probability distribution function  $p_i := Pr(x_i)$  for  $i = 1, \dots, n$ , then the Shannon entropy is defined as

$$H(X) := \sum_{i=1}^n p_i \log p_i.$$

Let  $X$  be a discrete random variable on a finite set  $\mathfrak{X}$  with probability distribution function  $P_X$ , let  $Y$  be a discrete random variable on a finite set  $\mathfrak{Y}$  with probability distribution function  $P_Y$ , and let  $P_{XY}$  be their joint probability distribution, then the mutual information between  $X$  and  $Y$  is defined as

$$I(X, Y) := \sum_{x \in \mathfrak{X}, y \in \mathfrak{Y}} P_{XY}(x, y) \log \frac{P_{XY}(x, y)}{P_X(x)P_Y(y)}.$$

Let  $N(x|x^n)$  be the number of occurrences of the symbol  $x$  in the sequence  $X^n$ . For a probability distribution  $P \in P(A)$  and  $\delta \geq 0$  let typical sequences and conditional typical sequence be defined as :

$$\mathcal{T}_P^n := \{x^n \in A^n : N(x|x^n) = nP(x) \forall x \in A\},$$

$$\mathcal{T}_{P,\delta}^n := \{x^n \in A^n : |N(x|x^n) - nP(x)| \leq \delta \sqrt{nP(x)(1-P(x))} \forall x \in A\}.$$

An  $(n, J_n)$  code for the compound wiretap channel  $(V_t, W_t)_{t \in \theta}$  consists of stochastic encoders  $\{E\} : \{1, \dots, J_n\} \rightarrow P(A^n)$  and a collection of mutually disjoint sets  $\{D_j \subset B^n : j \in \{1, \dots, J_n\}\}$  (decoding sets).

A non-negative number  $R$  is an achievable secrecy rate for the compound wiretap channel  $(W_t, V_t)$  in the case with CSI if there is a collection of  $(n, J_n)$  codes  $\{E_t : t \in \theta\}, \{D_j : j = 1, \dots, J_n\}$  such that

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log J_n \geq R,$$

$$\lim_{n \rightarrow \infty} \max_{t \in \theta} \max_{j \in \{1, \dots, J_n\}} \sum_{x^n \in A^n} E_t(x^n | j) W_t^n(D_j^c | x^n) = 0, \quad (1)$$

$$\lim_{n \rightarrow \infty} \max_{t \in \theta} I(J; Z_t^n) = 0, \quad (2)$$

where  $J$  is an uniformly distributed random variable with value in  $\{1, \dots, J_n\}$ , and  $Z_t^n$  are the resulting random variables at the output of wiretap channels  $V_t^n$ .

A non-negative number  $R$  is an achievable secrecy rate for the compound wiretap channel  $(W_t, V_t)$  in the case without CSI if there is a collection of  $(n, J_n)$  codes  $(E, \{D_j : j = 1, \dots, J_n\})$  such that

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log J_n \geq R,$$

$$\lim_{n \rightarrow \infty} \max_{t \in \theta} \max_{j \in \{1, \dots, J_n\}} \sum_{x^n \in A^n} E(x^n | j) W_t^n(D_j^c | x^n) = 0, \quad (3)$$

$$\lim_{n \rightarrow \infty} \max_{t \in \theta} I(J; Z_t^n) = 0. \quad (4)$$

*Remark 1:* A weaker and widely used security criterion is obtained if we replace (2), respectively (4), with  $\lim_{n \rightarrow \infty} \max_{t \in \theta} \frac{1}{n} I(J; Z_t^n) = 0$ .

In case with CSI, let  $p'_t(x^n) := \begin{cases} \frac{p_t^n(x^n)}{p_t^n(\mathcal{T}_{p_t,\delta}^n)} & \text{if } x^n \in \mathcal{T}_{p_t,\delta}^n \\ 0 & \text{else} \end{cases}$

and  $X^{(t)} := \{X_{j,l}^{(t)}\}_{j \in \{1, \dots, J_n\}, l \in \{1, \dots, L_{n,t}\}}$  be a family of random matrices whose entries are i.i.d. according to  $p'_t$ .

It was shown in [5] that for any  $\omega > 0$ , if we set

$$J_n = \lfloor 2^{n(\min_{t \in \theta} (I(p_t, V_t) - \frac{1}{n} \log L_{n,t}) - \mu)} \rfloor,$$

where  $\mu$  is a positive constant which does not depend on  $j, t$ , and can be arbitrarily small when  $\omega$  goes to 0, then there are such  $\{D_j : j = 1, \dots, J_n\}$  that for all  $t \in \theta$

$$\Pr \left( \sum_{j=1}^{J_n} \frac{1}{J_n} \sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} W_t^n(D_j^c | X_{j,l}^{(t)}) > \sqrt{T} 2^{-n\omega/2} \right) \leq \sqrt{T} 2^{-n\omega/2}. \quad (5)$$

Since here only the error of the legitimate receiver is analyzed, so for the result above just the channels  $V_t$ , but not those of the wiretapper, are regarded.

In view of (5), one has (see [5])

the largest achievable rate, called capacity, of the compound wiretap channel with CSI at the transmitter  $C_{S,CSI}$ , is given by

$$C_{S,CSI} = \min_{t \in \theta} \max_{V \rightarrow A \rightarrow (BZ)_t} (I(V, B_t) - I(V, Z_t)) , \quad (6)$$

where  $B_t$  are the resulting random variables at the output of legal receiver channels.  $Z_t$  are the resulting random variables at the output of wiretap channels.

Analogously, in case without CSI, the idea is similar to the case with CSI: Let  $p'(x^n) := \begin{cases} \frac{p^n(x^n)}{p^n(\mathcal{T}_{p,\delta}^n)} & \text{if } x^n \in \mathcal{T}_{p,\delta}^n \\ 0 & \text{else} \end{cases}$

and  $X^n := \{X_{j,l}\}_{j \in \{1, \dots, J_n\}, l \in \{1, \dots, L_n\}}$  be a family of random matrices whose components are i.i.d. according to  $p'$ .

For any  $\omega > 0$ , define

$$J_n = \lfloor 2^{n(\min_{t \in \theta} (I(p_t, V_t) - \frac{1}{n} \log L_n) - \mu)} \rfloor,$$

where  $\mu$  is a positive constant which does not depend on  $j, t$ , and can be arbitrarily small when  $\omega$  goes to 0, then there are such  $\{D_j : j = 1, \dots, J_n\}$  that for all  $t \in \theta$

$$\Pr \left( \sum_{j=1}^{J_n} \frac{1}{J_n} \sum_{l=1}^{L_n} \frac{1}{L_n} W_t^n(D_j(X)^c | X_{j,l}) > \sqrt{T} 2^{-n\omega/2} \right) \leq \sqrt{T} 2^{-n\omega/2}. \quad (7)$$

Using (7) one can obtain (see [5]) that the secrecy capacity of the compound wiretap channel without CSI at the transmitter  $C_S$  is lower bounded as follows,

$$C_S \geq \max_{V \rightarrow A \rightarrow (BZ)_t} (\min_{t \in \theta} I(V, B_t) - \max_{t \in \theta} I(V, Z_t)) . \quad (8)$$

### III. CLASSICAL COMPOUND QUANTUM WIRETAP CHANNELS

Let  $A$  and  $B$  be finite sets, and let  $H$  be a finite-dimensional complex Hilbert space. Let  $P(A)$  and  $P(B)$  be the sets of probability distributions on  $A$  and  $B$  respectively, and  $\mathcal{S}(H)$  be the space of self-adjoint, positive-semidefinite bounded linear operators with trace 1 on  $H$ . Let  $\theta := \{1, \dots, T\}$  and for every  $t \in \theta$  let  $W_t$  be a channel  $A \rightarrow P(B)$  and  $V_t$  be a classical quantum channel, i.e., a map  $A \rightarrow \mathcal{S}(H)$ :  $A \ni x \rightarrow V_t(x) \in \mathcal{S}(H)$ . We define  $(V_t, W_t)_{t \in \theta}$  as a classical compound quantum wiretap channel. Associate to  $V_t$  is the channel map on  $n$ -block  $V_t^{\otimes n}$ :  $A^n \rightarrow \mathcal{S}(H^{\otimes n})$  with  $V_t^{\otimes n}(x^n) := V_t(x_1) \otimes \dots \otimes V_t(x_n)$ .

For a state  $\rho$ , the von Neumann entropy is defined as

$$S(\rho) := -\text{tr}(\rho \log \rho).$$

Let  $P$  be a probability distribution over a finite set  $J$ , and  $\Phi := \{\rho(x) : x \in J\}$  be a set of states labeled by elements of  $J$ . Then the Holevo  $\chi$  quantity is defined as

$$\chi(P, \Phi) := S \left( \sum_{x \in J} P(x) \rho(x) \right) - \sum_{x \in J} P(x) S(\rho(x)).$$

An  $(n, J_n)$  code for the classical compound quantum wiretap channel  $(V_t, W_t)_{t \in \theta}$  consists of stochastic encoders  $\{E\} : \{1, \dots, J_n\} \rightarrow P(A^n)$  and a collection of mutually disjoint sets  $\{D_j \subset B^n : j \in \{1, \dots, J_n\}\}$  (decoding sets).

A non-negative number  $R$  is an achievable secrecy rate for the classical compound quantum wiretap channel  $(W_t, V_t)_{t \in \theta}$  with CSI if there is an  $(n, J_n)$  code  $(\{E_t : t \in \theta\}, \{D_j : j = 1, \dots, J_n\})$  such that

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{1}{n} \log J_n &\geq R, \\ \lim_{n \rightarrow \infty} \max_{t \in \theta} \max_{j \in \{1, \dots, J_n\}} \sum_{x^n \in A^n} E_t(x^n | j) W_t^n(D_j^c | x^n) &= 0, \\ \lim_{n \rightarrow \infty} \max_{t \in \theta} \chi(J; Z_t^{\otimes n}) &= 0, \end{aligned}$$

where  $J$  is a uniformly distributed random variable with value in  $\{1, \dots, J_n\}$ .  $Z_t$  are the sets of states such that the wiretapper will get.

A non-negative number  $R$  is an achievable secrecy rate for the classical compound quantum wiretap channel  $(W_t, V_t)_{t \in \theta}$  without CSI if there is an  $(n, J_n)$  code  $(E, \{D_j : j = 1, \dots, J_n\})$  such that

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{1}{n} \log J_n &\geq R, \\ \lim_{n \rightarrow \infty} \max_{t \in \theta} \max_{j \in \{1, \dots, J_n\}} \sum_{x^n \in A^n} E(x^n | j) W_t^n(D_j^c | x^n) &= 0, \\ \lim_{n \rightarrow \infty} \max_{t \in \theta} \chi(J; Z_t^{\otimes n}) &= 0. \end{aligned}$$

**Theorem 1:** The largest achievable rate (secrecy capacity) of the classical compound quantum wiretap channel  $(W_t, V_t)_{t \in \theta}$  in the case with CSI  $C_{S, CSI}$  at the transmitter is given by

$$C_{S, CSI} = \min_{t \in \theta} \max_{P \rightarrow A \rightarrow B_t Z_t} (I(P, B_t) - \chi(P, Z_t)). \quad (9)$$

Respectively, in the case without CSI, the secrecy capacity of the classical compound quantum wiretap channel  $(W_t, V_t)_{t \in \theta}$   $C_S$  is lower bounded as follows

$$C_S \geq \max_{P \rightarrow A \rightarrow B_t Z_t} (\min_{t \in \theta} I(P, B_t) - \max_t \chi(P, Z_t)), \quad (10)$$

where  $B_t$  are the resulting random variables at the output of legal receiver channels, and  $Z_t$  are the resulting random states at the output of wiretap channels.

*Proof: 1) Lower bound*

Let  $p'_t$ ,  $X^{(t)}$ , and  $D_j$  be defined like in classical case. Then (5) still holds since the sender transmits through a classical channel to the legitimate receiver. We abbreviate  $\mathcal{X} := \{X^{(t)} : t \in \theta\}$ .

(Analogously, in the case without CSI, let  $p'$ ,  $X^n$  and  $D_j$  be defined like in classical case, then (7) still holds.)

For  $\rho \in \mathcal{S}(H)$  and  $\alpha > 0$  there exists an orthogonal subspace projector  $\Pi_{\rho, \alpha}$  commuting with  $\rho^{\otimes n}$  and satisfying

$$\text{tr}(\rho^{\otimes n} \Pi_{\rho, \alpha}) \geq 1 - \frac{d}{\alpha^2}, \quad (11)$$

$$\text{tr}(\Pi_{\rho, \alpha}) \leq 2^{nS(\rho) + Kd\alpha\sqrt{n}}, \quad (12)$$

$$\Pi_{\rho, \alpha} \cdot \rho^{\otimes n} \cdot \Pi_{\rho, \alpha} \leq 2^{-nS(\rho) + Kd\alpha\sqrt{n}} \Pi_{\rho, \alpha}, \quad (13)$$

where  $a := \#\{A\}$ , and  $K$  is a constant which is in polynomial order of  $n$ .

For  $P \in P(A)$ ,  $\alpha > 0$  and  $x^n \in \mathcal{T}_P^n$  there exists an orthogonal subspace projector  $\Pi_{V, \alpha}(x^n)$  commuting with  $V_{x^n}^{\otimes n}$  and satisfying:

$$\text{tr}(V^{\otimes n}(x^n) \Pi_{V, \alpha}(x^n)) \geq 1 - \frac{ad}{\alpha^2}, \quad (14)$$

$$\text{tr}(\Pi_{V, \alpha}(x^n)) \leq 2^{nS(V|P) + Kd\alpha\sqrt{n}}, \quad (15)$$

$$\begin{aligned} \Pi_{V, \alpha}(x^n) \cdot V^{\otimes n}(x^n) \cdot \Pi_{V, \alpha}(x^n) \\ \leq 2^{-nS(V|P) + Kd\alpha\sqrt{n}} \Pi_{V, \alpha}(x^n), \end{aligned} \quad (16)$$

$$\text{tr}(V^{\otimes n}(x^n) \cdot \Pi_{PV, \alpha\sqrt{a}}) \geq 1 - \frac{ad}{\alpha^2}, \quad (17)$$

where  $a := \#\{A\}$ ,  $d := \dim H$ , and  $K$  is a constant which is in polynomial order of  $n$  (see [13]).

Let

$$Q_t(x^n) := \Pi_{PV_t, \alpha\sqrt{a}} \Pi_{V_t, \alpha}(x^n) \cdot V_t^{\otimes n}(x^n) \cdot \Pi_{V_t, \alpha}(x^n) \Pi_{PV_t, \alpha\sqrt{a}}$$

where  $\alpha$  will be defined later.

**Lemma 1** (see [14]): Let  $\rho$  be a state and  $X$  be a positive operator with  $X \leq id$  (the identity matrix) and  $1 - \text{tr}(\rho X) \leq \lambda \leq 1$ . Then

$$\|\rho - \sqrt{X} \rho \sqrt{X}\| \leq \sqrt{8\lambda}. \quad (18)$$

With the Lemma 1, (11), (17), and the fact that  $\Pi_{PV_t, \alpha\sqrt{a}}$  and  $\Pi_{V_t, \alpha}(x^n)$  are both projection matrices, for any  $t$  and  $x^n$  it holds:

$$\|Q_t(x^n) - V_t^{\otimes n}(x^n)\| \leq \frac{\sqrt{8(ad + d)}}{\alpha}. \quad (19)$$

We set  $\Theta_t := \sum_{x^n \in \mathcal{T}_{p'_t, \delta}^n} p'_t(x^n) Q_t(x^n)$ . For given  $z^n$  and  $t$ ,  $\langle z^n | \Theta_t | z^n \rangle$  is the expected value of  $\langle z^n | Q_t(x^n) | z^n \rangle$  under the condition  $x^n \in \mathcal{T}_{p'_t, \delta}^n$ .

**Lemma 2** (see [3]): Let  $\mathcal{V}$  be a finite dimensional Hilbert space,  $X_1, \dots, X_L$  be a sequence of i.i.d. random variables with values in  $\mathcal{S}(\mathcal{V})$  such that  $X_i \leq \mu \cdot id_{\mathcal{V}}$  for all  $i \in \{1, \dots, L\}$ , and  $\epsilon \in ]0, 1[$ . Let  $p$  be a probability distribution on  $\{X_1, \dots, X_L\}$ ,  $\rho = \sum_i p(X_i) X_i$  be the expected value of  $X_i$ , and  $\Pi'_{\rho, \lambda}$  be the projector onto the subspace spanned by the eigenvectors of  $\rho$  whose corresponding eigenvalues are greater than  $\frac{\lambda}{\dim \mathcal{V}}$ , then

$$\begin{aligned} \Pr \left( \left\| L^{-1} \sum_{i=1}^L X_i - \Pi'_{\rho, \lambda} \cdot \rho \cdot \Pi'_{\rho, \lambda} \right\| > \epsilon \right) \\ \leq 2 \cdot (\dim \mathcal{V}) \exp \left( -L \frac{\epsilon^2 \lambda}{2 \ln 2 (\dim \mathcal{V}) \mu} \right). \end{aligned} \quad (20)$$

Let  $\mathcal{V}$  be the image of  $\Pi_{P,\alpha\sqrt{a}}$ . By (12), we have

$$\dim \mathcal{V} \leq 2^{nS(P)+Kd\alpha\sqrt{an}}.$$

Furthermore

$$\begin{aligned} Q_t(x^n) &= \Pi_{P_{V_t},\alpha\sqrt{a}} \Pi_{V_t,\alpha}(x^n) \cdot V_t^{\otimes n}(x^n) \cdot \Pi_{V_t,\alpha}(x^n) \Pi_{P_{V_t},\alpha\sqrt{a}} \\ &\leq 2^{-n(S(V_t|P)+Kd\alpha\sqrt{n})} \Pi_{P_{V_t},\alpha\sqrt{a}} \Pi_{V_t,\alpha}(x^n) \Pi_{P_{V_t},\alpha\sqrt{a}} \\ &\leq 2^{-n \cdot S(V_t|P)+Kd\alpha\sqrt{n}} \cdot \Pi_{P_{V_t},\alpha\sqrt{a}} \\ &\leq 2^{-n \cdot S(V_t|P)+Kd\alpha\sqrt{n}} \cdot id_{\mathcal{V}}. \end{aligned} \quad (21)$$

The first inequality follows from (16). The second inequality holds because  $\Pi_{V_t,\alpha}$  and  $\Pi_{P_{V_t},\alpha\sqrt{a}}$  are projection matrices. The third inequality holds because  $\Pi_{P_{V_t},\alpha\sqrt{a}}$  is a projection matrix onto  $\mathcal{V}$ .

Thus, by (20) and (21)

$$\begin{aligned} &Pr \left( \left\| \sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} Q_t(X_{j,l}^{(t)}) - \Pi'_{\Theta_t,\lambda} \Theta_t \Pi'_{\Theta_t,\lambda} \right\| > \frac{1}{2}\epsilon \right) \\ &\leq 2 \cdot 2^{n(S(P)+Kd\alpha\sqrt{an})} \\ &\cdot \exp \left( -L_{n,t} \frac{\epsilon^2}{8 \ln 2} \lambda \cdot 2^{n(S(V_t|P)-S(P))+Kd\alpha\sqrt{n}(\sqrt{a}-1)} \right) \\ &= 2 \cdot 2^{n(S(P)+Kd\alpha\sqrt{an})} \\ &\cdot \exp \left( -L_{n,t} \frac{\epsilon^2}{8 \ln 2} \lambda \cdot 2^{n(-\chi(P,Z_t))+Kd\alpha\sqrt{n}(\sqrt{a}-1)} \right). \end{aligned}$$

the equality in the last line holds since

$$\begin{aligned} &S(P) - S(V_t|P) \\ &= S \left( \sum_j P(j) \sum_l \frac{1}{L_{n,t}} V_t^{\otimes n}(X_{j,l}^{(t)}) \right) \\ &- \sum_j P(j) S \left( \sum_l \frac{1}{L_{n,t}} V_t^{\otimes n}(X_{j,l}^{(t)}) \right) \\ &= \chi(P, Z_t). \end{aligned}$$

Notice that  $\|\Theta_t - \Pi'_{\Theta_t,\lambda} \Theta_t \Pi'_{\Theta_t,\lambda}\| \leq \lambda$ . Let  $\lambda := \frac{1}{2}\epsilon$  and  $n$  large enough then

$$\begin{aligned} &Pr \left( \left\| \sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} Q_t(X_{j,l}^{(t)}) - \Theta_t \right\| > \epsilon \right) \\ &\leq 2 \cdot 2^{n(S(P)+Kd\alpha\sqrt{an})} \\ &\cdot \exp \left( -L_{n,t} \frac{\epsilon^3}{16 \ln 2} \cdot 2^{n(-\chi(P,Z_t))+Kd\alpha\sqrt{n}(\sqrt{a}-1)} \right) \\ &\leq \exp \left( -L_{n,t} \cdot 2^{-n(\chi(P,Z_t)+\zeta)} \right), \end{aligned} \quad (22)$$

where  $\zeta$  is some suitable positive constant, which does not depend on  $j$ ,  $t$ , and can be arbitrarily small when  $\epsilon$  goes to 0.

Let  $L_{n,t} = 2^{n(\chi(P,Z_t)+2\zeta)}$  and  $n$  be large enough, then by (22) for all  $j$  it holds

$$Pr \left( \left\| \sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} Q_t(X_{j,l}^{(t)}) - \Theta_t \right\| > \epsilon \right) \leq \exp(-2^{n\zeta}) \quad (23)$$

and

$$\begin{aligned} &Pr \left( \left\| \sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} Q_t(X_{j,l}^{(t)}) - \Theta_t \right\| > \epsilon \forall t \right) \\ &= 1 - Pr \left( \bigcup_t \left\{ \left\| \sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} Q_t(X_{j,l}^{(t)}) - \Theta_t \right\| > \epsilon \right\} \right) \\ &\geq 1 - T \exp(-2^{n\zeta}) \\ &\geq 1 - 2^{-nv}, \end{aligned} \quad (24)$$

where  $v$  is some positive suitable constant which does not depend on  $j$  and  $t$ .

(Analogously, in the case without CSI, let  $L_n = 2^{n \max_t (\chi(P,Z_t)+\delta)}$  and  $n$  be large enough, then we can find some positive constant  $v$  so that

$$Pr \left( \left\| \sum_{l=1}^{L_n} \frac{1}{L_n} Q_t(X_{j,l}^{(t)}) - \Theta_t \right\| > \epsilon \forall t \right) \geq 1 - 2^{-nv} \quad (25)$$

for all  $j$ .)

*Remark 2:* Since  $\exp(-2^{n\zeta})$  converges to zero double exponentially faster, the inequality (24) remains true even if  $T$  depends on  $n$  and is exponentially large over  $n$ , i.e., we can still achieve exponentially small error.

From (5) and (24), it follows: For any  $\epsilon > 0$ , if  $n$  is large enough then the event

$$\begin{aligned} &\left( \bigcap_t \left\{ \sum_{j=1}^{J_n} \frac{1}{J_n} \sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} W_t^n(D_j^c(\mathcal{X})|X_{j,l}^{(t)}) \leq \epsilon \right\} \right) \\ &\cap \left( \bigcap_j \left\{ \left\| \sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} Q_t(X_{j,l}^{(t)}) - \Theta_t \right\| \leq \epsilon \forall t \right\} \right) \end{aligned}$$

has a positive probability. This means that we can find a realization  $x_{j,l}^{(t)}$  of  $X_{j,l}^{(t)}$  with a positive probability such that for all  $t \in \theta$ , we have

$$\sum_{j=1}^{J_n} \frac{1}{J_n} \sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} W_t^n(D_j^c|x_{j,l}^{(t)}) \leq \epsilon,$$

and

$$\left\| \sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} Q_t(x_{j,l}^{(t)}) - \Theta_t \right\| \leq \epsilon \forall j.$$

For any  $\gamma > 0$  let

$$R := \min_{t \in \theta} \max_{P \rightarrow A \rightarrow B_t Z_t} (I(P, B_t) - \chi(P, Z_t)) + \gamma,$$

then we have

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log J_n \geq R, \quad (26)$$

$$\lim_{n \rightarrow \infty} \max_{t \in \theta} \max_{j \in \{1, \dots, J_n\}} \sum_{x^n \in A^n} E_t(x^n | j) W_t^n(D_j^c | x^n) = 0, \quad (27)$$

where  $E_t$  is the random output of  $(X_{j,l}^{(t)})_l$ .

Choose a sufficiently large but fixed  $\alpha$  in (19) so that for all  $j$  it holds  $\|V_t^{\otimes n}(x_{j,l}^{(t)}) - Q_t(x_{j,l}^{(t)})\| < \epsilon$ . In this case, for any given  $j' \in \{1, \dots, J_n\}$  we have

$$\begin{aligned} & \left\| \sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} V_t^{\otimes n}(x_{j',l}^{(t)}) - \Theta_t \right\| \\ & \leq \left\| \sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} V_t^{\otimes n}(x_{j',l}^{(t)}) - \sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} Q_t(x_{j',l}^{(t)}) \right\| \\ & + \left\| \sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} Q_t(x_{j',l}^{(t)}) - \Theta_t \right\| \\ & \leq \sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} \|V_t^{\otimes n}(x_{j',l}^{(t)}) - Q_t(x_{j',l}^{(t)})\| \\ & + \left\| \sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} Q_t(x_{j',l}^{(t)}) - \Theta_t \right\| \\ & \leq 2\epsilon \end{aligned} \quad (28)$$

and  $\|\mathbb{E}_j \sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} V_t^{\otimes n}(x_{j,l}^{(t)}) - \Theta_t\| \leq \epsilon$  for any probability distribution uniformly distributed on  $\{1, \dots, J_n\}$ .

**Lemma 3 (Fannes inequality [14]):** Let  $\mathfrak{X}$  and  $\mathfrak{Y}$  be two states in a  $d$ -dimensional complex Hilbert space and  $\|\mathfrak{X} - \mathfrak{Y}\| \leq \mu < \frac{1}{e}$ , then

$$|S(\mathfrak{X}) - S(\mathfrak{Y})| \leq \mu \log d - \mu \log \mu. \quad (29)$$

If  $J$  is a probability distribution uniformly distributed on  $\{1, \dots, J_n\}$ , then from the inequality (28) and Lemma 3 we

have

$$\begin{aligned} & \chi(J; Z_t^{\otimes n}) \\ & = S \left( \mathbb{E}_j \sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} V_t^{\otimes n}(x_{j,l}^{(t)}) \right) \\ & - \sum_{j=1}^{J_n} J(j) S \left( \sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} V_t^{\otimes n}(x_{j,l}^{(t)}) \right) \\ & \leq |S \left( \mathbb{E}_j \sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} V_t^{\otimes n}(x_{j,l}^{(t)}) \right) - S(\Theta_t)| \\ & + |S(\Theta_t) - \sum_{j=1}^{J_n} J(j) S \left( \sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} V_t^{\otimes n}(x_{j,l}^{(t)}) \right)| \\ & \leq \epsilon \log d - \epsilon \log \epsilon \\ & + \left| \sum_{j=1}^{J_n} J(j) \left[ S(\Theta_t) - S \left( \sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} V_t^{\otimes n}(x_{j,l}^{(t)}) \right) \right] \right| \\ & \leq 3\epsilon \log d - \epsilon \log \epsilon - 2\epsilon \log 2\epsilon. \end{aligned} \quad (30)$$

We have

$$\lim_{n \rightarrow \infty} \max_{t \in \theta} \chi(J; Z_t^{\otimes n}) = 0. \quad (31)$$

(Analogously, in the case without CSI, we can find a realization  $x_{j,l}^n$  of  $X_{j,l}^{(t)}$  with a positive probability such that: For all  $t \in \theta$ , we have

$$\begin{aligned} & \sum_{j=1}^{J_n} \frac{1}{J_n} \sum_{l=1}^{L_n} \frac{1}{L_n} W_t^n(D_j^c | x_{j,l}) \leq \epsilon, \\ & \left\| \sum_{l=1}^{L_n} \frac{1}{L_n} Q_t(x_{j,l}) - \Theta_t \right\| \leq \epsilon \quad \forall j. \end{aligned}$$

For any  $\gamma > 0$  let

$$R := \max_{P \rightarrow A \rightarrow B_t Z_t} \left( \min_{t \in \theta} I(P, B_t) - \max_t \chi(P, Z_t) \right) + \gamma,$$

then we have

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log J_n \geq R, \quad (32)$$

$$\lim_{n \rightarrow \infty} \max_{t \in \theta} \max_{j \in \{1, \dots, J_n\}} \sum_{x^n \in A^n} E(x^n | j) W_t^n(D_j^c | x^n) = 0. \quad (33)$$

From  $\|\sum_{l=1}^{L_n} \frac{1}{L_n} V_t^{\otimes n}(x_{j,l}^{(t)}) - \Theta_t\| \rightarrow 0$  for  $n \rightarrow \infty$  it follows

$$\lim_{n \rightarrow \infty} \max_{t \in \theta} \chi(J; Z_t^{\otimes n}) = 0, \quad (34)$$

for any probability distribution  $J$  uniformly distributed on  $\{1, \dots, J_n\}$  in the case without CSI.)

Combining (5) and (31) (respectively (34)) we obtain

$$C_{S,CSI} \geq \min_{t \in \theta} \max_{V \rightarrow A \rightarrow B_t Z_t} (I(V, B_t) - \chi(V, Z_t)),$$

respectively

$$C_S \geq \max_{P \rightarrow A \rightarrow B_t Z_t} (\min_{t \in \theta} I(P, B_t) - \max_{t \in \theta} \chi(P, Z_t)) .$$

## 2) Upper bound for case with CSI

Considering  $(\mathcal{C}_n)$  is a sequence of  $(n, J_n)$  code such that

$$\sup_{t \in \theta} \frac{1}{J_n} \sum_{j=1}^{J_n} \sum_{x^n \in A^n} E(x^n | j) W_t^n(D_j^c | x^n) =: \epsilon_{1,n} , \quad (35)$$

$$\sup_{t \in \theta} \chi(J; Z_t^{\otimes n}) =: \epsilon_{2,n} , \quad (36)$$

where  $\lim_{n \rightarrow \infty} \epsilon_{1,n} = 0$  and  $\lim_{n \rightarrow \infty} \epsilon_{2,n} = 0$ ,  $J$  denotes the random variable which is uniformly distributed on the message set  $\{1, \dots, J_n\}$ .

Let  $C(V_t, W_t)$  denote the secrecy capacity of the wiretap channel  $(V_t, W_t)$  in the sense of [13]. Choose  $t' \in \theta$  such that  $C(V_{t'}, W_{t'}) = \min_{t \in \theta} C(V_t, W_t)$ .

It is well-known, in information theory, that even in the case without wiretapper (we have only one classical channel  $W_{t'}$ ), the capacity cannot exceed  $I(J; B_{t'}) + \xi$  for any constant  $\xi > 0$ . So the capacity of a quantum wiretap channel  $(V_{t'}, W_{t'})$  cannot be greater than

$$\begin{aligned} & I(J; B_{t'}) + \xi \\ & \leq \lim_{n \rightarrow \infty} [I(J; B_{t'}) - \chi(J; Z_{t'}^{\otimes n})] + \xi + \epsilon_{2,n} \\ & \leq [I(J; B_{t'}) - \chi(J; Z_{t'})] + \epsilon \end{aligned}$$

for any  $\epsilon > 0$ .

Since we cannot exceed the secrecy capacity of the worst wiretap channel, we have

$$C_{S,CSI} \leq \min_{t \in \theta} \max_{V \rightarrow A \rightarrow B_t Z_t} (I(V, B_t) - \chi(V, Z_t)) . \quad (37)$$

■

## IV. CLASSICAL QUANTUM COMPOUND WIRETAP CHANNEL WITH CSI

Let  $H$  be a finite-dimensional complex Hilbert space. Let  $\mathcal{S}(H)$  be the space of self-adjoint, positive-semidefinite bounded linear operators on  $H$  with trace 1. For every  $t \in \theta$  let  $W_t$  respectively  $V_t$  be quantum channels, i.e., completely positive trace preserving maps  $\mathcal{S}(H) \rightarrow \mathcal{S}(H)$ .

An  $(n, J_n, \lambda)$  code for the classical quantum compound wiretap channel  $(W_t, V_t)_{t \in \theta}$  consists of a family of vectors  $w := \{w(j) : j = 1, \dots, J_n\} \subset \mathcal{S}(H^{\otimes n})$  and a collection of positive semi-definite operators  $\{D_j : j \in \{1, \dots, J_n\}\} \subset \mathcal{S}(H^{\otimes n})$  which is a partition of the identity, i.e.  $\sum_{j=1}^{J_n} D_j = id_{H^{\otimes n}}$ .

A non-negative number  $R$  is an achievable secrecy rate for the classical quantum compound wiretap channel  $(W_t, V_t)_{t \in \theta}$  with CSI if there is an  $(n, J_n, \lambda)$  code  $(\{w_t := \{w_t(j) : j : t\}, \{D_j : j\})$  such that

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log J_n \geq R ,$$

$$\begin{aligned} \lim_{n \rightarrow \infty} \max_{t \in \theta} \frac{1}{J_n} \sum_{j=1}^{J_n} \text{tr}(W_t^{\otimes n}(w_t(j)) D_j) & \geq 1 - \lambda , \\ \lim_{n \rightarrow \infty} \max_{t \in \theta} \chi(J; Z_t^{\otimes n}) & = 0 , \end{aligned}$$

where  $J$  is a uniformly distributed random variable with value in  $\{1, \dots, J_n\}$ , and  $Z_t$  are the sets of states such that the wiretapper will get.

*Theorem 2:* The largest achievable rate (secrecy capacity) of the classical quantum compound wiretap channel in the case with CSI is given by

$$C_{CSI} = \lim_{n \rightarrow \infty} \min_{t \in \theta} \max_{P, w_t} \frac{1}{n} (\chi(P, B_t^{\otimes n}) - \chi(P, Z_t^{\otimes n})) , \quad (38)$$

where  $B_t$  are the resulting random states at the output of legal receiver channels, and  $Z_t$  are the resulting random states at the output of wiretap channels.

*Proof:* Our idea is to send the information in two parts, firstly, we send the state information with finite blocks of finite bits with a code  $C_1$  to the receiver, and then, depending on  $t$ , we send the message with a code  $C_2^{(t)}$  in the second part.

### 1) Sending channel state information with finite bits

For the first part, we don't require that the first part should be secure against the wiretapper, since we assume that the wiretapper already has the full knowledge of the CSI.

By ignoring the security against the wiretapper, we have only to look at the compound channel  $(W_t)_{t \in \theta}$ . Let  $W = (W_t)_t$  be an arbitrary compound classical quantum channel. Then by [4], for each  $\lambda \in (0, 1)$  the  $\lambda$ -capacity  $C(W, \lambda)$  equals

$$C(W, \lambda) = \inf_t \max_p \chi(p, W_t) . \quad (39)$$

If  $\min_t \max_p \chi(p, W_t) > 0$  holds, then the sender can build a code  $C_1$  such that the CSI can be sent to the legal receiver with a block with length  $l \leq \frac{\log T}{\min_t \max_p \chi(p, W_t)}$ . We need to do nothing because in this case the right hand side of (38) is zero.

Let  $c = 1 - \lambda$ , then for any required upper bound  $\delta = 2^{-c'}$ , with given  $c' > 0$ , the sender can repeat sending this block  $\log c \cdot c'$  times, and the legal receiver simply picks out the state that he receives most frequently to find out  $t$  with a error probability  $\leq \delta$ .

The first part is of length  $l \cdot \log c \cdot c' = O(1)$ , which is negligible compared to the second part.

### 2) Message transformation when both the sender and the legal receiver know CSI

If both the sender and the legal receiver have the full knowledge of  $t$ , then we only have to look at the single wiretap channel  $(W_t, V_t)$ .

In [7] and [8], it is shown that there exists an  $(n, J_n, \lambda)$  code for the quantum wiretap channel  $(W, V)$  with

$$\log J_n = \max_{P, w} (\chi(P, B^{\otimes n}) - \chi(P, Z^{\otimes n})) - \epsilon , \quad (40)$$

for any  $\epsilon > 0$ , where  $B$  is the resulting random variable at the output of legal receiver's channel and  $Z$  the output of the wiretap channel.

When the sender and the legal receiver both know  $t$ , they can build an  $(n, J_{n,t}, \lambda)$  code  $C_2^{(t)}$  where

$$\log J_{n,t} = \max_{P, w_t} (\chi(V, B_t^{\otimes n}) - \chi(V, Z_t^{\otimes n})) - \epsilon. \quad (41)$$

Thus,

$$C_{CSI} \geq \lim_{n \rightarrow \infty} \min_{t \in \theta} \max_{P, w_t} \frac{1}{n} (\chi(P, B_t^{\otimes n}) - \chi(P, Z_t^{\otimes n})). \quad (42)$$

*Remark 3:* For the construction of the second part of our code, we use random coding and request that the randomization can be sent (see [7]). However, it is shown in [5] that the randomization could not always be sent if we require that we use one unique code which is secure against the wiretapper and suitable for every channel state, i.e., it does not depend on  $t$ . This is not a counterexample to our results above, neither to the construction of  $C_1$  nor to the construction of  $C_2^{(t)}$ , because of following facts.

The first part of our code does not need to be secure. For our second part, the legal transmitters can use the following strategy: At first they build a code  $C_1 = (E, \{D_j : j = 1, \dots, J_n\})$  and a code  $C_2^{(t)} = (E^{(t)}, \{D_j^{(t)} : j = 1, \dots, J_n\})$  for every  $t \in \theta$ . If the sender wants to send the CSI  $t' \in \theta$  and the message  $j$ , he encodes  $t'$  with  $E$  and  $j$  with  $E^{(t')}$ , then he sends both parts together through the channel. After receiving both parts, the legal receiver decodes the first part with  $\{D_j : j\}$ , and chooses the right decoders  $\{D_j^{(t')} : j\} \in \{\{D_j^{(t)} : j\} : t \in \theta\}$  to decode the second part. With this strategy, we can avoid using one unique code which is suitable for every channel state.

### 3) Upper bound

For any  $\epsilon > 0$  choose  $t' \in \theta$  such that  $C(V_{t'}, W_{t'}) \leq \inf_{t \in \theta} C(V_t, W_t) + \epsilon$ .

From [7] and [8], we know that the capacity of the quantum wiretap channel  $(W_{t'}, V_{t'})$  cannot be greater than

$$\lim_{n \rightarrow \infty} \max_{P, w_{t'}} \frac{1}{n} (\chi(P, B_{t'}^{\otimes n}) - \chi(P, Z_{t'}^{\otimes n})).$$

Since we cannot exceed the capacity of the worst wiretap channel, we have

$$C_{CSI} \leq \lim_{n \rightarrow \infty} \min_{t \in \theta} \max_{P, w_t} \frac{1}{n} (\chi(P, B_t^{\otimes n}) - \chi(P, Z_t^{\otimes n})). \quad (43)$$

This together with (42) completes the proof of Theorem 2. ■

*Remark 4:* In [12], it is shown that if for a given  $t$  and any  $n \in \mathbb{N}$

$$I(P, B_t^{\otimes n}) \geq I(P, Z_t^{\otimes n})$$

holds for all  $P \in P(A)$  and  $\{w_t(j) : j = 1, \dots, J_n\} \subset S(H^{\otimes n})$ , then

$$\begin{aligned} & \lim_{n \rightarrow \infty} \max_{P, w_t} \frac{1}{n} (\chi(P, B_t^{\otimes n}) - \chi(P, Z_t^{\otimes n})) \\ &= \max_{P, w_t} (\chi(P, B_t) - \chi(P, Z_t)). \end{aligned}$$

Thus if for every  $t \in \theta$  and  $n \in \mathbb{N}$ ,

$$I(P, B_t^{\otimes n}) \geq I(P, Z_t^{\otimes n})$$

holds for all  $P \in P(A)$  and  $\{w_t(j) : j = 1, \dots, J_n\} \subset S(H^{\otimes n})$ , we have

$$C_{CSI} = \min_{t \in \theta} \max_{P, w_t} (\chi(P, B_t) - \chi(P, Z_t)).$$

So far, we assumed that  $|\theta|$ , the number of the channels, is  $< \infty$ . Now we look at the case where  $|\theta|$  can be arbitrary.

*Theorem 3:* For an arbitrary set  $\theta$  we have

$$C_{CSI} = \lim_{n \rightarrow \infty} \inf_{t \in \theta} \max_{P, w_t} \frac{1}{n} (\chi(P, B_t^{\otimes n}) - \chi(P, Z_t^{\otimes n})). \quad (44)$$

*Proof:* Let  $W : \mathcal{S}(H) \rightarrow \mathcal{S}(H)$  be a linear map, then let

$$\|W\|_{\diamond} := \sup_{n \in \mathbb{N}} \max_{a \in S(\mathbb{C}^n \otimes H), \|a\|_1 = 1} \|(id_n \otimes W)(a)\|_1 \quad (45)$$

where  $\|\cdot\|_1$  stands for the trace norm.

It is well known [11] that this norm is multiplicative, i.e.  $\|W \otimes W'\|_{\diamond} = \|W\|_{\diamond} \cdot \|W'\|_{\diamond}$ .

A  $\tau$ -net in the space of the completely positive trace preserving maps is a finite set  $(W^{(k)})_{k=1}^K$  with the property that for each  $W$  there is at least one  $k \in \{1, \dots, K\}$  with  $\|W - W^{(k)}\|_{\diamond} < \tau$ .

*Lemma 4 ( $\tau$ -net [10]):* For any  $\tau \in (0, 1]$  there is a  $\tau$ -net of quantum-channels  $(W_t^{(k)})_{k=1}^K$  in the space of the completely positive trace preserving maps with  $K \leq (\frac{3}{\tau})^{2d^4}$ , where  $d = \dim H$ .

If  $|\theta|$  is arbitrary, then for any  $\xi > 0$  let  $\tau = \frac{\xi}{-\log \xi}$ . By Lemma 4 there exists a finite set  $\theta'$  with  $|\theta'| \leq (\frac{3}{\tau})^{2d^4}$  and  $\tau$ -nets  $(W_{t'})_{t' \in \theta'}$ ,  $(V_{t'})_{t' \in \theta'}$  such that for every  $t \in \theta$  we can find a  $t' \in \theta'$  with  $\|W_t - W_{t'}\|_{\diamond} \leq \tau$  and  $\|V_t - V_{t'}\|_{\diamond} \leq \tau$ . For every  $t' \in \theta'$  the legal transmitters build a code  $C_2^{(t')} = \{w_{t'}, \{D_{t',j} : j\}\}$ . Since by [7], the error of the code  $C_2^{(t')}$  decreases exponentially to its length, we can find an  $N = O(-\log \xi)$  such that for all  $t' \in \theta'$  it holds

$$\frac{1}{J_N} \sum_{j=1}^{J_N} \text{tr}(W_{t'}^{\otimes N}(w_{t'}(j)) D_{t',j}) \geq 1 - \lambda - \xi, \quad (46)$$

$$\chi(J; Z_{t'}^{\otimes N}) \leq \xi, \quad (47)$$

Then, if the sender obtains the state information “ $t$ ”, he can send with finite bits “ $t'$ ” to the legal receiver in the first part, and then they build a code  $C_2^{(t')}$  that fulfills (46) and (47) to transmit the message.

For every  $t'$  and  $j$  let  $\psi_{t'}(j) \in H^{\otimes n} \otimes H^{\otimes n}$  be an arbitrary purification of the state  $w_{t'}(j)$ . Then we have

$$\begin{aligned}
& \text{tr} [(W_t^{\otimes N} - W_{t'}^{\otimes N})(w_{t'}(j))] \\
&= \text{tr} (\text{tr}_{H^{\otimes N}} [id_H^{\otimes N} \otimes (W_t^{\otimes N} - W_{t'}^{\otimes N})(|\psi_{t'}(j)\rangle\langle\psi_{t'}(j)|)]) \\
&= \text{tr} [id_H^{\otimes N} \otimes (W_t^{\otimes N} - W_{t'}^{\otimes N})(|\psi_{t'}(j)\rangle\langle\psi_{t'}(j)|)] \\
&= \|id_H^{\otimes N} \otimes (W_t^{\otimes N} - W_{t'}^{\otimes N})(|\psi_{t'}(j)\rangle\langle\psi_{t'}(j)|)\|_1 \\
&\leq \|W_t^{\otimes N} - W_{t'}^{\otimes N}\|_{\diamond} \cdot \|(|\psi_{t'}(j)\rangle\langle\psi_{t'}(j)|)\|_1 \\
&\leq N\tau.
\end{aligned}$$

The first equality follows from the definition of purification. the second equality follows from the definition of trace. The third equality follows from the fact that  $\|A\|_1 = \text{tr}(A)$  for any self-adjoint, positive-semidefinite bounded linear operator  $A$ . The first inequality follows by the definition of  $\|\cdot\|_{\diamond}$ . The second inequality follows from the facts that  $\|(|\psi_{t'}(j)\rangle\langle\psi_{t'}(j)|)\|_1 = 1$  and  $\|W_t^{\otimes N} - W_{t'}^{\otimes N}\|_{\diamond} = \|(W_t - W_{t'})^{\otimes N}\|_{\diamond} = N \cdot \|W_t - W_{t'}\|_{\diamond}$ , since  $\|\cdot\|_{\diamond}$  is multiplicative.

It follows

$$\begin{aligned}
& \frac{1}{J_N} \sum_{j=1}^{J_N} \text{tr} (W_t^{\otimes N}(w_{t'}(j)) D_{t',j}) \\
& - \frac{1}{J_N} \sum_{j=1}^{J_N} \text{tr} (W_{t'}^{\otimes N}(w_{t'}(j)) D_{t',j}) \\
&= \frac{1}{J_N} \sum_{j=1}^{J_N} \text{tr} [(W_t^{\otimes N} - W_{t'}^{\otimes N})(w_{t'}(j)) D_{t',j}] \\
&\leq \frac{1}{J_N} \sum_{j=1}^{J_N} \text{tr} [(W_t^{\otimes N} - W_{t'}^{\otimes N})(w_{t'}(j))] \\
&\leq \frac{1}{J_N} J_N N \cdot \tau \\
&= N\tau.
\end{aligned} \tag{48}$$

$N\tau$  tends to zero when  $\xi$  goes to zero, since  $N = O(-\log \xi)$ .

Let  $J$  be a probability distribution uniformly distributed on  $\{1, \dots, J_N\}$ , and  $\{\rho(j) : j = 1, \dots, J_N\}$  be a set of states labeled by elements of  $J$ . By Lemma 3 we have

$$\begin{aligned}
& \|\chi(J, V_t) - \chi(J, V_{t'})\| \\
&\leq \|S \left( \sum_{j=1}^{J_N} J(j) V_t(\rho(j)) \right) - S \left( \sum_{j=1}^{J_N} J(j) V_{t'}(\rho(j)) \right)\| \\
&+ \left\| \sum_{j=1}^{J_N} J(j) S(V_t(\rho(j))) - \sum_{j=1}^{J_N} J(j) S(V_{t'}(\rho(j))) \right\| \\
&\leq 2\tau \log d - 2\tau \log \tau,
\end{aligned} \tag{49}$$

since by  $\|V_t - V_{t'}\|_{\diamond} \leq \tau$ , it holds  $\|V_t(\rho) - V_{t'}(\rho)\| \leq \tau$  for all  $\rho \in \mathcal{S}(H)$ .

By (48) and (49) it holds

$$\begin{aligned}
& \max_t \frac{1}{J_N} \sum_{j=1}^{J_N} \text{tr} (W_t^{\otimes N}(w_{t'}(j)) D_{t',j}) \geq 1 - \lambda - \xi - N\tau, \\
& \chi(J; Z_t^{\otimes N}) \leq \xi + 2\tau \log d - 2\tau \log \tau.
\end{aligned}$$

Since  $N\tau$  and  $2\tau \log d$  both tend to zero when  $\xi$  goes to zero, we have

$$\begin{aligned}
& \lim_{n \rightarrow \infty} \max_{t \in \theta} \frac{1}{J_n} \sum_{j=1}^{J_n} \text{tr} (W_t^{\otimes n}(w_{t'}(j)) D_{t',j}) \geq 1 - \lambda, \\
& \lim_{n \rightarrow \infty} \chi(J; Z_t^{\otimes n}) = 0.
\end{aligned}$$

The bits that the sender uses to transform the CSI is large but constant, so it is still negligible compared to the second part. We obtain

$$C_{CSI} > \lim_{n \rightarrow \infty} \inf_{t \in \theta} \max_{P, w_t} \frac{1}{n} (\chi(P, B_t^{\otimes n}) - \chi(P, Z_t^{\otimes n})). \tag{50}$$

The proof of the converse is similar to those given in the proof of Theorem 2, where we consider a worst  $t'$ . ■

*Remark 5:* For Theorem 2 and Theorem 3, we have only required that the probability that the legal receiver does not obtain the correct message tends to zero when the code length goes to infinity. We have not specified how fast it should tends to zero. If we analyze the relation between the error probability  $\varepsilon$  and the code length, then we have the following facts.

In the case of finite  $\theta$ , let  $\varepsilon_1$  denote the probability that the legal receiver does not obtain the correct CSI, and let  $\varepsilon_2$  denote the probability that the legal receiver, having CSI, does not obtain the correct message. Since the length of first part of the code is  $l \cdot \log c \cdot c' = O(\log \varepsilon_1)$ , as we defined in Section IV, we have  $\varepsilon_1^{-1}$  is  $O(\exp(l \cdot \log c \cdot c')) = O(\exp(n))$ , where  $n$  stands for the length of first part. And for the second part of the code,  $\varepsilon_2$  decreased exponentially to the length of the second part, as proven in [7]. Thus, the error probability  $\varepsilon = \max\{\varepsilon_1, \varepsilon_2\}$  decreases exponentially to the code length in the case of finite  $\theta$ .

If  $\theta$  is infinite, let  $\varepsilon_1$  denote the probability that the legal receiver does not obtain the correct CSI. Then we have to build two  $\tau$ -nets, each contains  $O((\frac{-\log \varepsilon_1}{\varepsilon_1})^{-2d^4})$  channels. If we want to send the CSI of these  $\tau$ -nets,  $l$ , as defined in Section IV, will be  $O(-2d^4 \cdot \log(\varepsilon_1 \log \varepsilon_1))$ , this means here  $\varepsilon_1^{-1}$  will be  $O(\exp(\frac{n}{4d^4})) = O(\exp(n))$ , where  $n$  stands for the length of first part. So we can still achieve that the error probability decreases exponentially to the code length in case of infinite  $\theta$ .

#### ACKNOWLEDGMENT

We thank Igor Bjelakovic and Holger Boche for useful discussions. Support by the Bundesministerium für Bildung und Forschung (BMBF) via grant 01BQ1052 is gratefully acknowledged.



## REFERENCES

- [1] R. Ahlswede, I. Bjelakovic, H. Boche, and J. Nötzel, Quantum capacity under adversarial quantum noise: arbitrarily varying quantum channels submitted to Communications in Mathematical Physics.
- [2] R. Ahlswede and N. Cai, Transmission, identification and common randomness capacities for wire-tape channels with secure feedback from the decoder, General Theory of Information Transfer and Combinatorics, Lecture Notes in Computer Science, Vol. 4123, Springer Verlag, 258-275, 2006.
- [3] R. Ahlswede and A. Winter, Strong converse for identification via quantum channels, IEEE Trans. Inform. Theory, Vol. 48, No. 3, 569-579, 2002. Addendum: IEEE Trans. Inform. Theory, Vol. 49, No. 1, 346, 2003.
- [4] I. Bjelakovic and H. Boche, Classical capacities of averaged and compound quantum channels. IEEE Trans. Inform. Theory, Vol. 57, No. 7, 3360-3374, 2009.
- [5] I. Bjelakovic, H. Boche, and J. Sommerfeld, Capacity results for compound wiretap channels, CoRR abs, 1103-2013, 2011.
- [6] D. Blackwell, L. Breiman, and A. J. Thomasian, The capacity of a class of channels, Ann. Math. Stat. Vol. 30, No. 4, 1229-1241, 1959.
- [7] N. Cai, A. Winter, and R. W. Yeung, Quantum privacy and quantum wiretap channels, Problems of Information Transmission, Vol. 40, No. 4, 318-336, 2004.
- [8] I. Devetak, The private classical information capacity and quantum information capacity of a quantum channel, IEEE Trans. Inform. Theory, Vol. 51, No. 1, 44-55, 2005.
- [9] Y. Liang, G. Kramer, H. Poor, and S. Shamai, Compound wiretap channels, EURASIP Journal on Wireless Communications and Networking, Article ID 142374, 2008.
- [10] V. D. Milman and G. Schechtman, Asymptotic Theory of Finite Dimensional Normed Spaces. Lecture Notes in Mathematics 1200, Springer-Verlag, corrected second printing, Berlin, 2001.
- [11] V. Paulsen, Completely Bounded Maps and Operator Algebras, Cambridge Studies in Advanced Mathematics 78, Cambridge University Press, Cambridge, UK, 2002.
- [12] S. Watanabe, Remarks on Private and Quantum Capacities of More Capable and Less Noisy Quantum Channels, arXiv:1110-5746 Vol. [quant-ph], 2011.
- [13] M. Wilde, From Classical to Quantum Shannon Theory, arXiv:1106-1445, 2011.
- [14] A. Winter: Coding theorem and strong converse for quantum channels, IEEE Trans. Inform. Theory, Vol. 45, No. 7, 2481-2485, 1999.
- [15] A. D. Wyner, The wire-tap channel, Bell System Technical Journal, Vol. 54, No. 8, 1355-1387, 1975.